

Predictors of Earnings Risk with Machine Learning

Ethan Ballou*

Scott Drewianka*

January 12, 2026

Abstract

This paper looks at the determinants of lifetime earnings risk under a Restricted Income Profile (RIP) model using traditional and machine learning methods such as lasso and SHAP values. The paper builds on the work of Drewianka and Oberg (2025) which uses a moment condition approach derive a parameter that captures permanent income risk. The paper finds that education and age are important in explaining lifetime earnings risk. The paper also finds that macroeconomic variables such as probability of recession and real GDP growth are important and along with state controls may further imply a role of government policy. Finally, the paper finds that occupation controls are important while industry controls do not appear to play a strong role.

Keywords: machine learning, restricted income profile, earnings instability, risk

JEL Codes: D8, J0, D3

*University of Wisconsin - Milwaukee

1. NEED TO ADD IN SHAP RANK TABLES

- 1 Introduction
- 2 Data and Model
- 3 Empirical Strategy
- 4 Results
- 5 Conclusion

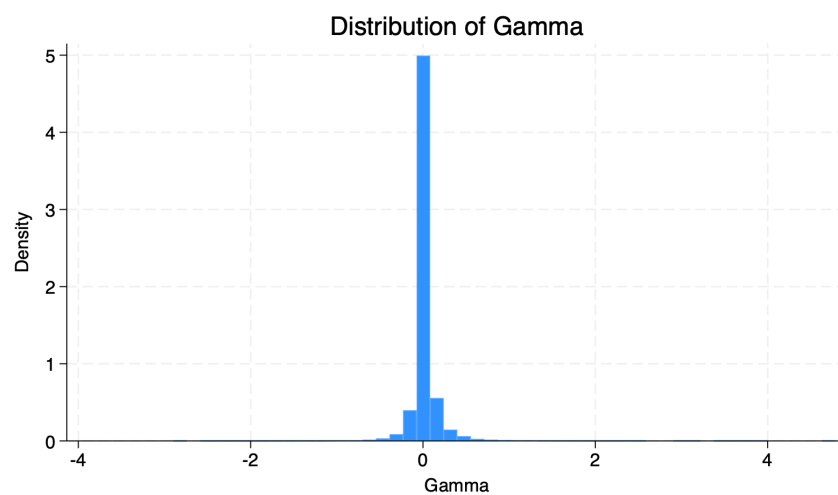


Figure 1: Distribution of Gamma

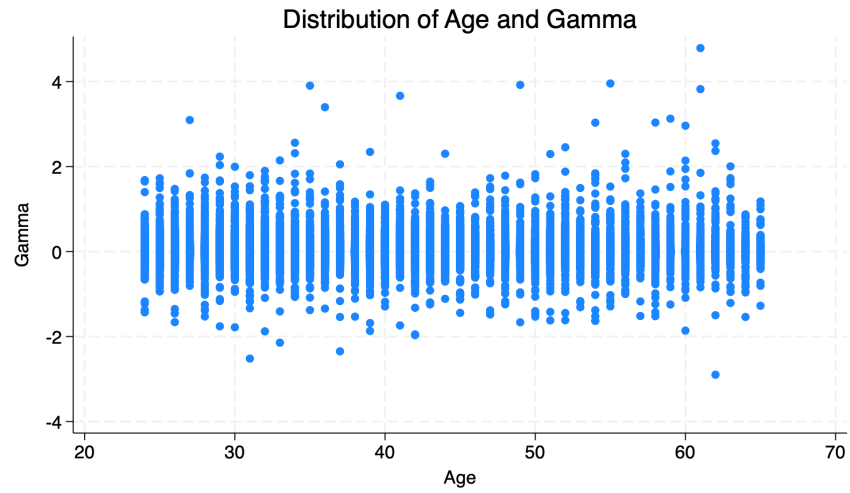


Figure 2: Scatterplot of Age vs. Gamma

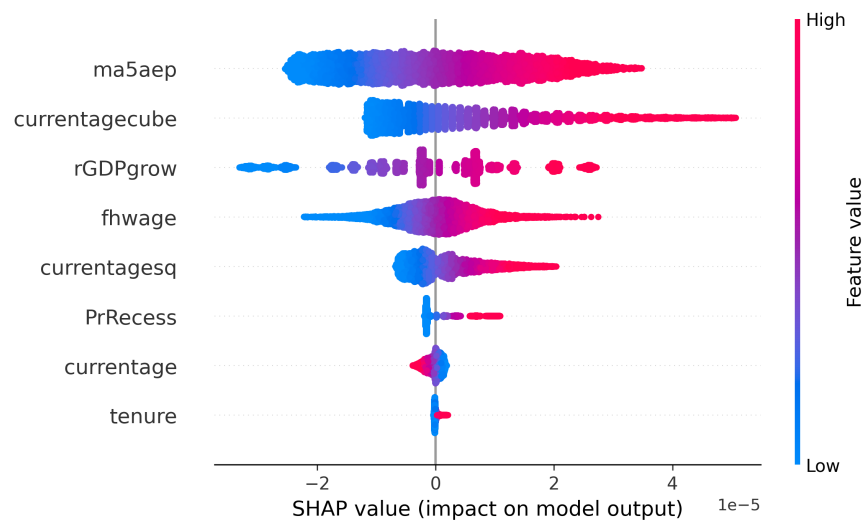


Figure 3: SHAP Summary Plot for Gamma

Table 1: Gamma Regressions: OLS Results

	(1)	(2)	(3)	(4)	(5)
	Gamma	Gamma	Gamma	Gamma	Gamma
Less than High School	-0.0115*** (0.00215)	-0.0107*** (0.00232)	-0.0117*** (0.00242)	-0.0109*** (0.00253)	-0.0117*** (0.00255)
High School Graduate	-0.00918*** (0.00144)	-0.00854*** (0.00149)	-0.00902*** (0.00158)	-0.00812*** (0.00171)	-0.00850*** (0.00173)
Some College	-0.00746*** (0.00170)	-0.00681*** (0.00172)	-0.00736*** (0.00177)	-0.00674*** (0.00184)	-0.00695*** (0.00185)
Probability of Recession	0.0000422 (0.0000437)	-0.0104 (0.0179)	-0.0000953 (51.52)	-0.0000817 (51.50)	-0.0000292 (51.49)
Real GDP growth rate	0.000600 (0.000400)	-0.00432 (0.00325)	-0.0000211 (3.981)	-0.000186 (3.979)	-0.000626 (3.979)
5-year moving average of AEP	0.0000636* (0.0000295)	0.0000629 (0.0000322)	0.0000374 (0.0000340)	0.0000766* (0.0000347)	0.0000688 (0.0000356)
Out of Labor Force	-0.00430 (0.00535)	-0.00410 (0.00536)	-0.00471 (0.00573)	-0.00461 (0.00572)	-0.00474 (0.00572)
Tenure	-0.0000438 (0.0000888)	-0.0000336 (0.0000963)	0.0000277 (0.0000981)	0.00000375 (0.0000984)	0.0000156 (0.0000987)
Age	0.00592* (0.00269)	0.00650* (0.00273)	0.00679* (0.00274)	0.00679* (0.00274)	0.00671* (0.00274)
Age Squared	-0.000170** (0.0000651)	-0.000193** (0.0000659)	-0.000196** (0.0000662)	-0.000197** (0.0000662)	-0.000194** (0.0000663)
Age Cubed	0.00000153** (0.000000509)	0.00000172*** (0.000000515)	0.00000173*** (0.000000517)	0.00000173*** (0.000000518)	0.00000171*** (0.000000518)
State FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Race FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes
Industry FE	No	No	Yes	No	Yes
R-squared	0.001	0.002	0.003	0.005	0.006
N	82357	82333	81556	81556	81556

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Gamma Regressions: Stepwise Selection Results

	(1)	(2)	(3)	(4)	(5)
	Gamma	Gamma	Gamma	Gamma	Gamma
Less than High School	-0.0110*** (0.00211)	-0.0110*** (0.00211)	-0.0101*** (0.00205)	-0.0108*** (0.00234)	-0.0110*** (0.00236)
High School Graduate	-0.00904*** (0.00143)	-0.00905*** (0.00143)	-0.00840*** (0.00143)	-0.00829*** (0.00166)	-0.00834*** (0.00168)
Some College	-0.00740*** (0.00169)	-0.00740*** (0.00169)	-0.00731*** (0.00172)	-0.00723*** (0.00182)	-0.00725*** (0.00183)
Age	0.00571* (0.00268)	0.00568* (0.00268)		0.00594* (0.00270)	0.00566* (0.00270)
Age Cubed	0.00000150** (0.000000508)	0.00000149** (0.000000508)	0.000000439*** (8.71e-08)	0.00000150** (0.000000511)	0.00000143** (0.000000511)
5-year moving average of AEP	0.0000579* (0.0000286)	0.0000577* (0.0000286)		0.0000788* (0.0000316)	0.0000732* (0.0000324)
Age Squared	-0.000165* (0.0000650)	-0.000165* (0.0000650)	-0.0000282*** (0.00000578)	-0.000168* (0.0000654)	-0.000161* (0.0000654)
Constant	-0.0391 (0.0358)	-0.0387 (0.0358)	0.0395*** (0.00358)	-0.0414 (0.0360)	-0.0380 (0.0360)
State FE					
Year FE					
Race FE					
Cohort FE					
Occupation FE				✓	✓
Industry FE			✓		✓
State FE Available	No	Yes	Yes	Yes	Yes
Year FE Available	No	Yes	Yes	Yes	Yes
Race FE Available	No	Yes	Yes	Yes	Yes
Cohort FE Available	No	Yes	Yes	Yes	Yes
Occupation FE Available	No	No	No	Yes	Yes
Industry FE Available	No	No	Yes	No	Yes
R-squared	0.001	0.001	0.002	0.003	0.004
N	82357	82333	81556	81556	81556

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Gamma Regressions: Lasso Selection Results

	No Controls	No Occ/Ind	No Occ	No Ind	All Controls
EDU1	2	3	3	2	2
EDU2	1	2	2	2	2
EDU3	3	4	4	4	3
OLF	9	9	8	7	6
PrRecess	10				
currentage	8	7	6	5	4
currentagecube	4	8	7	6	5
currentagesq	11	10	10	9	8
ma5aep	5	5	5	3	3
rGDPgrow	6				
tenure	7	6	9	8	7
State FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Race FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes
Industry FE	No	No	Yes	No	Yes

Table 4: Gamma Regressions: Lasso Industry Selection Results

Industry	Selection Order
	1
Other Services	1
Priv. Househld	1
Legal Services	2
Construction	3
Financial Inst	3
Public Administration	3
Clothing/Text	4
Health Service	4
Mechanical Eng	5
Energy/Water	6
Agric.,Forestry	7
Chemicals	7
Educ./Sport	7
Earth/Clay/Stone	8
Retail	8
Social Security	9
Volunt./Church	9
Electrical Eng	10
Iron/Steel	10
Wholesale	10
Constr. Relate	11
Mining	11
Synthetics	11
Fisheries	12
Wood/Paper/Print	13
Not Applicable	14
Other Trans.	15
Restaurants	16
Train System	16
Food Industry	17
Insurance	17
Service Indust	18
Postal System	

Table 5: Gamma Regressions: Lasso Occupation Selection Results

Occupation	Selection Order
Dr./Dentist/Vet	1
Insurance Rep.	2
Transp. Attend	2
	3
Aero/MarineEngr	4
Author	4
Priv.Bus.Leadr	4
Lumbrman/Axman	5
Hair Stylist	6
Painter	6
Vendor	6
Chemist	7
Janitor	7
Legislator	7
Buyer	8
Farm Manager	8
Music/Perform	8
Labor/Craftsmn	9
Jewelry Maker	10
Mathematician	11
Prof. Athlete	11
Tel. Operator	11
Miner	12
Educator	13
Machine Fitter	13
Stone Cutter	13
BusinessManagr	14
ComputerOperat	14
SecurityServic	14
Cook/Waiter	15
Fisher/Hunter	15
Not Applicable	15
Agriculturladm	16
ChemicalWorker	17
Conductor	17
Eng.Tech.Expert	17
Stenographer	18
Transport.Oper	18
Architect/Engineer	19
Economist	19
Tech.Salespers	19
Lawyer	20
Soldier	20
Office Manager	21
Salesperson	21
Administrator	22
Ofc.Worker Etc	22

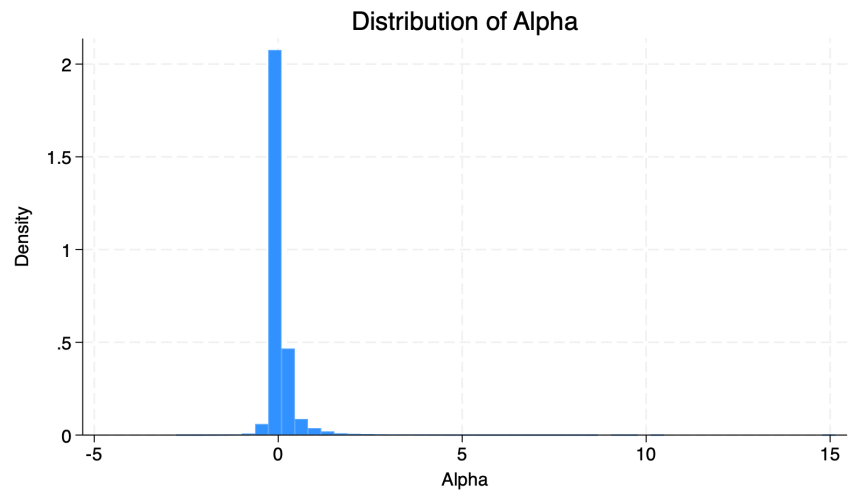


Figure 4: Distribution of Alpha

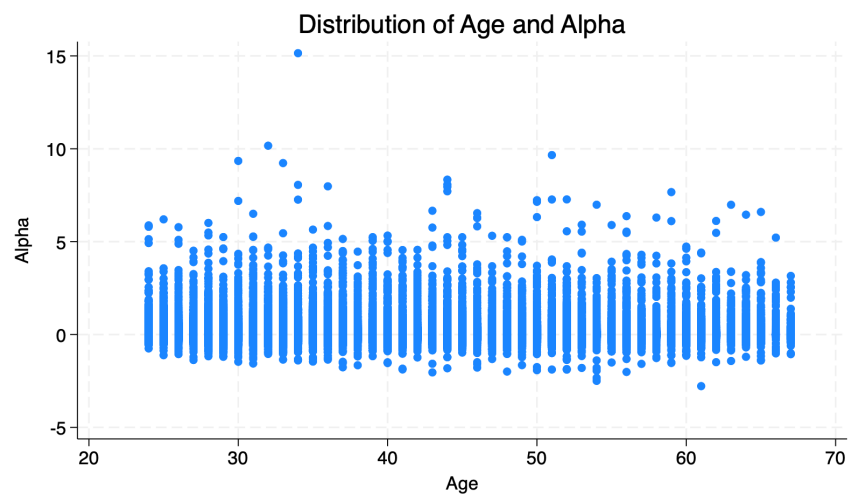


Figure 5: Scatterplot of Age vs. Alpha

Table 6: Alpha Regressions: OLS Results

	(1) Alpha	(2) Alpha	(3) Alpha	(4) Alpha	(5) Alpha
Less than High School	-0.0268*** (0.00451)	-0.0240*** (0.00487)	-0.0308*** (0.00508)	-0.0278*** (0.00531)	-0.0301*** (0.00535)
High School Graduate	-0.0131*** (0.00309)	-0.0129*** (0.00320)	-0.0136*** (0.00340)	-0.00909* (0.00368)	-0.0104** (0.00372)
Some College	-0.00240 (0.00366)	-0.00321 (0.00370)	-0.00223 (0.00384)	0.00135 (0.00398)	0.000562 (0.00401)
Probability of Recession	-0.0000420 (0.0000929)	-0.138* (0.0554)	-0.113 (0.0723)	-0.112 (0.0723)	-0.113 (0.0722)
Real GDP growth rate	-0.00144 (0.000855)	0.000541 (0.00633)	-0.00107 (0.00713)	-0.00287 (0.00710)	-0.00359 (0.00714)
5-year moving average of AEP	0.00127*** (0.0000616)	0.00138*** (0.0000669)	0.00103*** (0.0000712)	0.00111*** (0.0000726)	0.00106*** (0.0000745)
Out of Labor Force	0.113*** (0.0101)	0.111*** (0.0101)	0.0518*** (0.0111)	0.0521*** (0.0111)	0.0517*** (0.0111)
Tenure	-0.000698*** (0.000185)	-0.00000130 (0.000202)	0.000135 (0.000207)	0.0000259 (0.000208)	0.0000861 (0.000208)
Age	0.0165** (0.00520)	0.0183*** (0.00526)	0.0175*** (0.00530)	0.0182*** (0.00530)	0.0173** (0.00530)
Age Squared	-0.000398** (0.000124)	-0.000451*** (0.000125)	-0.000431*** (0.000126)	-0.000452*** (0.000126)	-0.000427*** (0.000126)
Age Cubed	0.00000324*** (0.000000947)	0.00000371*** (0.000000955)	0.00000356*** (0.000000963)	0.00000374*** (0.000000963)	0.00000353*** (0.000000964)
State FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Race FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes
Industry FE	No	No	Yes	No	Yes
R-squared	0.008	0.013	0.020	0.021	0.023
N	102946	102910	100781	100781	100781

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Alpha Regressions: Stepwise Selection Results

	(1) Alpha	(2) Alpha	(3) Alpha	(4) Alpha	(5) Alpha
Less than High School	-0.0261*** (0.00418)	-0.0204*** (0.00450)	-0.0259*** (0.00464)	-0.0259*** (0.00477)	-0.0272*** (0.00480)
High School Graduate	-0.0122*** (0.00271)	-0.0108*** (0.00278)	-0.0107*** (0.00289)	-0.00852** (0.00304)	-0.00913** (0.00306)
Age Squared	-0.000401** (0.000124)	-0.000394** (0.000124)	-0.000336** (0.000125)	-0.000384** (0.000125)	-0.000351** (0.000125)
Age Cubed	0.00000326*** (0.000000947)	0.00000327*** (0.000000950)	0.00000287** (0.000000959)	0.00000321*** (0.000000959)	0.00000296** (0.000000959)
Real GDP growth rate	-0.00115* (0.000541)				
5-year moving average of AEP	0.00127*** (0.0000601)	0.00135*** (0.0000650)	0.00106*** (0.0000689)	0.00111*** (0.0000711)	0.00107*** (0.0000728)
Out of Labor Force	0.113*** (0.0101)	0.112*** (0.0101)	0.0535*** (0.0111)	0.0541*** (0.0111)	0.0535*** (0.0111)
Tenure	-0.000704*** (0.000185)				
Age	0.0167** (0.00519)	0.0159** (0.00523)	0.0132* (0.00527)	0.0153** (0.00527)	0.0140** (0.00527)
Constant	-0.190** (0.0704)	-0.157 (0.0825)	-0.102 (0.0831)	-0.117 (0.0831)	-0.105 (0.0831)
State FE		✓	✓	✓	✓
Year FE		✓	✓	✓	✓
Race FE		✓	✓	✓	✓
Cohort FE					
Occupation FE				✓	✓
Industry FE			✓		✓
State FE Available	No	Yes	Yes	Yes	Yes
Year FE Available	No	Yes	Yes	Yes	Yes
Race FE Available	No	Yes	Yes	Yes	Yes
Cohort FE Available	No	Yes	Yes	Yes	Yes
Occupation FE Available	No	No	No	Yes	Yes
Industry FE Available	No	No	Yes	No	Yes
R-squared	0.008	0.012	0.018	0.021	0.022
N	102946	102910	100781	100781	100781

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Alpha Regressions: Lasso Selection Results

	No Controls	No Occ/Ind	No Occ	No Ind	All Controls
EDU1	4	5	4	4	4
EDU2	5	6	6	7	7
EDU3	7	7	8	6	6
OLF	2	3	3	3	3
PrRecess	8				
currentage		9	9	8	9
currentagecube	3	4	5	5	5
currentagesq		10	10	9	10
ma5aep	1	2	2	2	2
rGDPgrow	6				
tenure	4	8	7	10	8
State FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Race FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	Yes
Industry FE	No	No	Yes	No	Yes

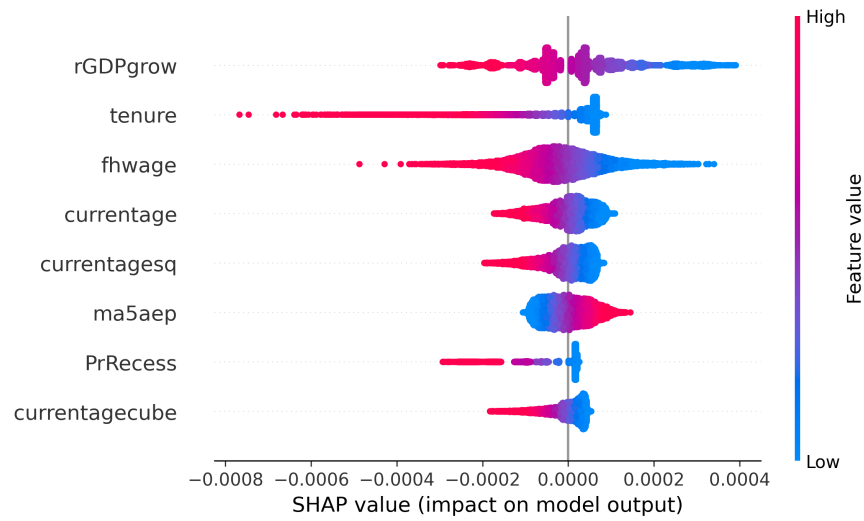


Figure 6: SHAP Summary Plot for Alpha

Table 9: Alpha Regressions: Lasso Industry Selection Results

Industry	Selection Order
Not Applicable	1
Agric.,Forestry	2
	3
Fisheries	3
Mechanical Eng	4
Public Administration	4
Construction	5
Legal Services	5
Other Services	6
Energy/Water	7
Constr. Relate	8
Clothing/Text	9
Synthetics	9
Service Indust	10
Iron/Steel	11
Postal System	12
Health Service	13
Chemicals	14
Food Industry	14
Train System	14
Educ./Sport	15
Electrical Eng	16
Retail	16
Other Trans.	17
Priv. Househld	18
Restaurants	18
Wood/Paper/Print	18
Financial Inst	19
Earth/Clay/Stone	20
Volunt./Church	21
Social Security	22
Insurance	23
Mining	
Wholesale	

Table 10: Alpha Regressions: Lasso Occupation Selection Results

Occupation	Selection Order
Not Applicable	1
Farm Manager	2
Fisher/Hunter	3
	4
Mathematician	5
Insurance Rep.	6
Architect/Engineer	7
Agriculturladm	8
Bricklay/Carpt	8
Inspector	8
Priv.Bus.Leadr	9
Prof. Athlete	9
Music/Perform	10
BusinessManagr	11
Mailman	11
ComputerOperat	12
Eng.Tech.Expert	12
Hair Stylist	12
Cook/Waiter	13
Ofc.Worker Etc	13
Scientist	14
Chemist	15
Economist	15
Tech.Salespers	15
Aero/MarineEngr	16
Office Manager	16
Sculptr/Paintr	16
Jewelry Maker	17
Printer Etc.	17
Rest./StoreMgr	17
Soldier	18
Administrator	19
Stenographer	19
Tool/Die Maker	19
Buyer	20
ChemicalWorker	20
Spinner/Weaver	20
Electr. Enginr	21
Machine Fitter	21
RelatMedicalJob	21
Convey. Oper.	22
Glazier	23
Miner	24
Lawyer	25
Shoemaker	25
Cleric	26
Dr./Dentist/Vet	26